

AD-A112 169

DECISION SCIENCE CONSORTIUM INC FALLS CHURCH VA  
RECONCILING CONTINUOUS PROBABILITY ASSESSMENTS.(U)  
OCT 81 D V LINDLEY

F/6 12/1

N00014-81-C-0330

UNCLASSIFIED

TR-81-5

.NL

1 of 1  
AD A  
112169

END  
DATE  
FILMED  
14-82  
DTIC

12



**DECISION SCIENCE CONSORTIUM, INC.**

AD A 1 1 2 1 6 9

RECONCILING CONTINUOUS PROBABILITY ASSESSMENTS

D. V. Lindley

Decision Science Consortium Inc.  
7700 Leesburg Pike, Suite 421  
Falls Church, Virginia 22043

October 1981

Prepared for:

OFFICE OF NAVAL RESEARCH  
Mathematical Sciences Division  
Naval Analysis Programs  
800 N. Quincy Street  
Arlington, Virginia 22217

MAR 19 1982

E

Technical Report No. 81-5

DTIC FILE COPY

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle)  RECONCILING CONTINUOUS PROBABILITY ASSESSMENTS		5. TYPE OF REPORT & PERIOD COVERED  Technical
7. AUTHOR(s)  D. V. Lindley		6. PERFORMING ORG. REPORT NUMBER 81-5
9. PERFORMING ORGANIZATION NAME AND ADDRESS DECISION SCIENCE CONSORTIUM, Inc. 7700 Leesburg Pike, Suite 421 Falls Church, Virginia 22043		8. CONTRACT OR GRANT NUMBER(s)  N00014-81-C-0330
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Analysis Prog. Mathematical and Information Sciences Division OFFICE OF NAVAL RESEARCH, Code 431 800 N. Quincy St., Arlington, VA 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61153N RR 014-11-01 NR 277-271
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE October 1981
		13. NUMBER OF PAGES 41
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for publication release: distribution unlimited. Reproduction in whole or in part is permitted for any purpose of the U.S. Government.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Decision Analysis Incoherence Probability Distributions Bayesian Statistics		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  Several subjects each provide a mean and a standard deviation for a single uncertain quantity, expressing thereby their own opinions of the value of the quantity. In the paper we discuss the problem of incorporating all these judgments into a single probability distribution for the quantity. Under simple, normal assumptions the resulting procedure is least squares. Dis- advantages of this are alleviated by using t-, instead of normal, distributions. Further refinements incorporate extra information provided by the standard deviations. These lead to nonnormal distributions and improvements to least		

DD FORM 1473  
1 JAN 73

EDITION OF 1 NOV 65 IS OBSOLETE  
S/N 0102-LF-014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Block 20.

Abstract (continued)

squares.

S/N 0102- LF- 014-6601

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

RECONCILING CONTINUOUS PROBABILITY ASSESSMENTS

By

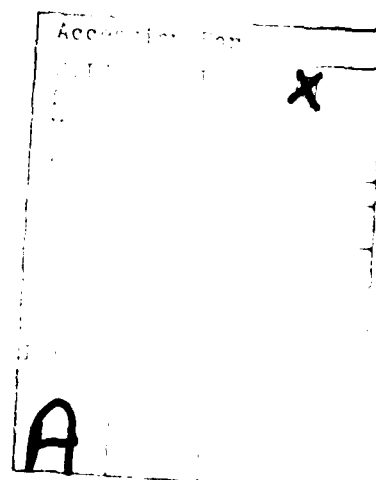
D.V. Lindley

Prepared for:

Office of Naval Research  
Mathematical Sciences Division  
Naval Analysis Programs  
Contract No. N00014-81-C-0330

October 1981

Decision Science Consortium, Inc.  
7700 Leesburg Pike, Suite 421  
Falls Church, Virginia 22043  
(703) 790-0510



## SUMMARY

Several subjects each provide a mean and a standard deviation for a single uncertain quantity, expressing thereby their own opinions of the value of the quantity. In the paper we discuss the problem of incorporating all these judgments into a single probability distribution for the quantity. Under simple, normal assumptions the resulting procedure is least squares. Disadvantages of this are alleviated by using  $t$ -, instead of normal, distributions. Further refinements incorporate extra information provided by the standard deviations. These lead to non-normal distributions and improvements to least squares.

## TABLE OF CONTENTS

1.0	INTRODUCTION-----	1
2.0	LIKELIHOOD PRINCIPLE-----	3
3.0	SINGLE SUBJECT, NO SCALE INFORMATION-----	6
3.1	Subject Trials-----	8
4.0	SEVERAL SUBJECTS, NO SCALE INFORMATION-----	10
4.1	Correlations Between Subjects-----	11
4.2	Example-----	12
5.0	t-DISTRIBUTION-----	16
6.0	SINGLE SUBJECT, WITH SCALE INFORMATION-----	20
6.1	Approximations-----	22
6.2	Example-----	23
7.0	SEVERAL SUBJECTS, SCALE INFORMATION-----	27
8.0	DISCUSSION-----	28
	REFERENCES-----	29
	APPENDIX A-----	A-1
A.1	Single Subject, No Scale Information-----	A-1
A.1.1	Subject trials-----	A-1
A.2	Several Subjects, No Scale Information-----	A-2
A.2.1	Correlations between subjects-----	A-2
A.3	Uncertainty About $\gamma$ : t-Distribution-----	A-3
A.4	Products of t-Likelihoods-----	A-4
A.5	Single Subject, Scale Information-----	A-5
A.5.1	Approximations-----	A-6
A.6	Other Forms of Scale Information-----	A-7
A.7	Several Subjects, Scale Information-----	A-8

## 1.0 INTRODUCTION

We are concerned in this paper with the assessment of an unknown, or uncertain, quantity  $\theta$ ; for example, the range of a target or the demand for a product. A subject  $S$  expresses his uncertainty about  $\theta$  in the form of a probability distribution for  $\theta$ , or at least provides some features of such a distribution: for example, its mean and variance. In the case of a target,  $S$  may be a sonarman who has used sonar devices to assess its position. With the demand for a product,  $S$  could be an experienced sales representative. The subject need not be a human being: thus the sonar device might directly produce a mean; or a market research organization might carry out a survey of the product's likely acceptability. We suppose there are several subjects,  $S_1, S_2, \dots, S_n$ , each producing his own assessment of  $\theta$ . In addition, there is an investigator  $N$  who receives these assessments and is required to provide an overall probability distribution for  $\theta$ . Thus  $N$  may be the ship's commander who, in addition to advice from the sonarman  $S_1$ , receives range estimates based on deflection/elevation angle  $S_2$ , and an Ekelund range  $S_3$ . In the product example,  $N$  may be the board of the company producing it, receiving advice on demand from several sources.

The standard approach used in problems of this type is based on least squares, in which the various estimates (or means) for  $\theta$  are combined linearly with weights that depend on the stated variances and possibly on any correlations

perceived between the subjects. In this paper some refinements of least squares are proposed that take into account more features of subjects' probability assessments.

The general principles behind this problem have been discussed by Lindley et al. (1979). The case of the subjectivist assessment of probabilities for events, rather than quantities, was considered by Lindley (1982) and French (1980). The least-squares approach has been used by Cohen and Brown (1980). Morris (1974, 1977) had important, pioneering papers on the subject. DeGroot (1980) has considered the important problem of how subjects might improve their probability judgments. Bordley and Wolff (1981) have discussed the problem in the light of the impossibility results of Dalkey.

Work on this paper was performed under Contract No. N00014-81-C-0330 for the Office of Naval Research, U.S. Navy, through Decision Science Consortium, Inc., Falls Church, Virginia. I am grateful to Rex V. Brown for many stimulating discussions on the topics of the paper and many related, practical issues.

## 2.0 LIKELIHOOD PRINCIPLE

We first discuss a general principle that is basic to all the analyses, beginning with the case of a single subject S. Even with a single subject S, N still has a problem because he may suspect S of biases, or overconfidence. It is therefore important to distinguish between S's probabilities for  $\theta$  and N's: they may well be different. We suppose S states certain features of his probability distribution for  $\theta$ . For example, he may state the three quartiles, saying that for him the probability is one quarter that  $\theta$  exceeds the upper quartile, and similarly for the others. In this paper we shall consider only the case where S states his mean,  $m$ , and standard deviation,  $s$ , for  $\theta$ . (All that is needed is that S provides measures of location and spread, referred to as  $m$  and  $s$ .) If in addition he says the distribution is normal, then his probabilities for  $\theta$  are completely specified, but we shall not assume this. Generally then S provides features  $t_1, t_2, \dots, t_m$  of his probability distribution for  $\theta$ .

In addition, N will have a distribution for  $\theta$  which we write  $p(\theta)$ . The notation  $p(\cdot)$  for probability will always refer to N's assessment, not S's. Having received S's information in the form of  $t_1, t_2, \dots, t_m$ , N requires to calculate  $p(\theta|t_1, t_2, \dots, t_m)$ , his revised distribution for  $\theta$  given that information. He should do this by Bayes theorem

$$p(\theta|t_1, t_2, \dots, t_m) \propto p(t_1, t_2, \dots, t_m|\theta)p(\theta). \quad (2.1)$$

In explanation, the first and last of the probabilities in the formula are the ones already introduced; namely N's assessment of  $\theta$  with and without S's information. The formula says the former is proportional to the product of the latter and a new probability,  $p(t_1, t_2, \dots, t_m | \theta)$ . This is N's probability that S will announce  $t_1, t_2, \dots, t_m$  when the value of the quantity is truly  $\theta$ . The technical name is likelihood for  $\theta$ . Thus N is effectively saying: were the range (or demand) truly to be  $\theta$ , what would I expect S to say; values with high probability are likely but those with small probability are unexpected. Notice that these are probabilities for N about what S says: they are not probabilities for S, which do not concern N except through the features,  $t_1, t_2, \dots, t_m$ , provided.

The general principle is that, in order to incorporate S's information into his assessment (that is, to change  $p(\theta)$  into  $p(\theta | t_1, t_2, \dots, t_m)$ ) N requires his likelihood for  $\theta$ : his assessment of what S will say were the true value to be  $\theta$ . Notice the inversion involved here: in order to make probability statements about  $\theta$ , N requires probabilities for the  $t$ 's, given  $\theta$ . Consider the case where S just states his mean for  $\theta$ : we write this as  $m$ , replacing the general notation  $t_1$ . Then N has to consider the probability that S will state  $m$  when the quantity is  $\theta$ :  $p(m | \theta)$ . If this distribution has mean  $\theta$ , that is, if  $E(m | \theta) = \theta$ , then N is saying that in his opinion (for these are all judgments by N, apart from the value of  $m$ ) S is free from bias.

If  $E(m | \theta) = \theta + 2$  then N expects S to overestimate by 2. The variance of  $p(m | \theta)$  describes N's appreciation of S's precision in assessing  $\theta$ ; a low value saying he is precise, a high value expressing little trust in S. We therefore

see that the likelihood reflects N's opinion about S as an evaluator of  $\theta$ . Readers familiar with Bayesian statistics will notice that N is treating  $t_1, t_2, \dots, t_m$  as data to be processed in the same way as any other data. The fact that the data reflect S's opinions does not affect the method of analysis, only the structure of the likelihood.

The principle extends to several subjects. If  $S_1$  and  $S_2$  both provide means for  $\theta$ ,  $m_1$  and  $m_2$ ; then N will require for insertion into Bayes formula,  $p(m_1, m_2 | \theta)$ , the probability that with range (or demand)  $\theta$ ,  $S_1$  will say  $m_1$  and  $S_2$ ,  $m_2$ . This incorporates not only the concepts of bias and precision already referred to but also any correlation between  $S_1$  and  $S_2$ . Thus if the subjects share much of their knowledge there may be positive correlation between them.

### 3.0 SINGLE SUBJECT, NO SCALE INFORMATION

Using the principle we study first the case of a single subject  $S$  who states his mean  $m$ , and standard deviation  $s$ , for  $\theta$ . Thus the sonarman may say that he thinks the target is at 10,000 meters with a standard deviation of 1,000 meters. This may be expressed indirectly by his saying that he is 95% sure that the target is between 8,000 and 12,000 meters away. With an implied normal distribution, these values are  $m \pm 2s$ . With  $m$  and  $s$  the features of the subject's distribution for  $\theta$ , the investigator has first to assess his probability that  $S$  will announce values  $m$  and  $s$ , given  $\theta$ :  $p(m, s|\theta)$ . This distribution can be factored

$$p(m, s|\theta) = p(m|s, \theta) p(s|\theta), \quad (3.1)$$

the probability  $p(s|\theta)$  that  $S$  will state  $s$  for the deviation, and  $p(m|s, \theta)$ , the probability that  $S$  will state  $m$ , given that he has stated  $s$ . We make

Assumption 1  $p(s|\theta)$  does not depend on  $\theta$ .

In other words, this says that  $N$  does not think that the distance the target is away will affect  $S$ 's perceived precision of his estimate: or, in the other example, the company thinks that the representative will think it just as easy to estimate a high demand as a low one. The assumption may well not be true:  $N$  may think  $S$  will perceive distant targets harder to range than near ones. We explore the relatively simple consequences of the assumption before proceeding in Section 5.0 to the more complicated case where it is relaxed.

For this case, Bayes formula (2.1), reads

$$\begin{aligned} p(\theta|m, s) &\propto p(m, s|\theta) p(\theta) \\ &= p(m|s, \theta) p(s|\theta) p(\theta), \end{aligned}$$

by the factorization, and if  $p(s|\theta)$  does not depend on  $\theta$  it may be absorbed into the constant of proportionality to give

$$p(\theta|m, s) \propto p(m|s, \theta) p(\theta). \quad (3.2)$$

Assumption 2.  $p(m|s, \theta)$  is normal, with mean  $\alpha + \beta\theta$  and standard deviation  $\gamma s$ .

To understand this, begin with the case  $\beta = 1$ . Then  $\alpha$  is a bias term. N is expressing the opinion that if the true value is  $\theta$ , he expects S to announce  $\theta + \alpha$ ; or, using the normality,  $\theta + \alpha$  is the most likely value for m. If  $\theta + \alpha$  is expected, how far away from it is S likely to be? This is described by  $\gamma$ . Using a 95% interval, N thinks that S could state values for m as much as  $2\gamma s$  from the expected value. If  $\gamma = 1$ , N is effectively agreeing with S's assessment. If  $\gamma > 1$ , N is inflating S's value, thinking that S is overconfident in his assessment. If  $\gamma < 1$ , N thinks S is lacking in confidence. For example, if the ship's commander felt that the sonar device did not take into account all the errors, he would use  $\gamma > 1$ . The company who felt the representative was unduly cautious might have  $\gamma < 1$ .

Thus  $\alpha$  and  $\gamma$  allow N to express his views about two major features of S's assessment as he, N, perceives it; namely the bias and the confidence. The remaining value  $\beta$  allows the bias to change with  $\theta$ . The bias is zero at  $\theta = \alpha/(1-\beta)$  and, if  $\beta > 1$ , increases with  $\theta$ . Thus  $\alpha = 0$ ,  $\beta = 1.1$  allows for a bias of 10% of the value of  $\theta$ ; overestimating the range by 10%. In much of the subsequent work we take  $\beta = 1$ , removing the bias dependence on  $\theta$ .

Note that although N judges that  $s$  is uninfluenced by  $\theta$  (assumption 1),  $s$  still plays an important role in N's final judgment of  $\theta$ . This is because he uses  $s$  (through  $\gamma s$ ) to say how reliable he thinks  $m$  is as an evaluation of  $\theta$ . Thus  $s$ , on its own, is uninformative, but, in conjunction with  $m$ , provides useful information.

If N has little other knowledge of  $\theta$  except that provided by S, it is easy to show (A.1)\* that with assumptions 1 and 2, N's distribution for  $\theta$  is also normal with mean  $(m-\alpha)/\beta$  and standard deviation  $\gamma\beta^{-1}s$ , specializing to  $(m-\alpha)$  and  $\gamma s$  when  $\beta = 1$ . In that case, all N does is to correct  $m$  for bias and adjust  $s$  for confidence. The case  $\alpha = 0$ ,  $\gamma = 1$  leads to N agreeing with S's stated values. N is sometimes said to think S is calibrated.

### 3.1 Subject Trials

We therefore see that in order to process S's information N needs two values,  $\alpha$  and  $\gamma$  (and also  $\beta$  if not 1). It is important to realize that

\*References beginning with A refer to sections in the Appendix to the paper. Equations therein, begin with a.

$\alpha$  and  $\gamma$  express N's opinion of S as an assessor: for example, the ship's commander's view of the sonar device. It is, therefore, not sensible to say that N does not know the values of  $\alpha$  and  $\gamma$ : for they are expressions of what he does know. Nevertheless, it does make sense to say that N could improve his knowledge of S, and hence his values of  $\alpha$  and  $\gamma$ , by studying S. For example, ship's trials may be performed in which S states  $m_i$  and  $s_i$  in trial  $i$  for a target whose true range (unknown to him) is  $\theta_i$ . With data  $(m_i, s_i, \theta_i)$  from  $n$  trials, reasonable values for  $\alpha$  and  $\gamma^2$  (see a.3) are  $\hat{\alpha} = \sum (m_i - \theta_i) s_i^{-2} / \sum s_i^{-2}$  the weighted average bias, and  $\hat{\gamma}^2 = \sum (m_i - \theta_i - \hat{\alpha})^2 / s_i^2 n$ .

Even this treatment may be unsatisfactory if it is felt that trials carried out under non-combative situations are not totally reliable as a guide to what might happen in action. Consider, for example, just the matter of confidence, expressed through  $\gamma$ . If N thinks that the sonar device is less reliable in action but that S will stick to his training scheme, N may well increase  $\gamma$  from  $\hat{\gamma}$ . On the contrary, if N thinks that the device retains its reliability but S loses confidence in the face of the enemy,  $\hat{\gamma}$  may need to be decreased. The point is that  $\alpha$ ,  $\beta$  and  $\gamma$  are judgmental. The judgmental element can be reduced by trials, or other data collection, but can rarely be eliminated.

An alternative treatment of uncertainty about  $\gamma$  that has certain advantages will be discussed in Section 5.0.

#### 4.0 SEVERAL SUBJECTS, NO SCALE INFORMATION

Consider next the case of several subjects  $S_1, S_2, \dots, S_n$  each providing means and standard deviations  $(m_i, s_i)$ , where N's opinions of each  $S_i$  satisfies assumptions 1 and 2, so that, for each  $S_i$ , N has a triplet  $(\alpha_i, \beta_i, \gamma_i)$ : thus N may have different opinions about the biases and confidences of the subjects. What is N's overall opinion of  $\theta$  in the light of all this information? It is not possible to answer this question without considering another aspect of the subject's assessments, namely their potential correlations. For example, if  $S_1$  overestimates  $\theta (m_1 > \theta)$ , will  $S_2$  tend to do the same? If this is not so; that is, if the judgments of the subjects are, in N's view, unrelated, we say N considers them to be independent. In that case the effect of each  $S_i$  is to multiply N's probabilities by the likelihood for  $S_i$ ,  $p(m_i | s_i, \theta)$ , as in equation (3.2); and the final result (a.4) is to take the individual means  $(m_i - \alpha_i)/\beta_i$  and calculate a weighted average of them, the weights being the inverses of the individual variances  $\gamma_i^2 \beta_i^{-2} s_i^2$ , termed the precisions. The final precision is the total of the individual precisions (a.5).

In the case where N considers each  $S_i$  to be calibrated ( $\alpha_i = 0, \beta_i = \gamma_i = 1$ ) this is the usual least squares procedure mentioned earlier. The general, uncalibrated case uses a modification of that procedure, adjusting the raw  $(m_i, s_i)$  before combining by least squares.

#### 4.1 Correlations Between Subjects

If N does not judge the subjects to be independent, then he has to make a judgment about the extent of the dependence. If he continues to think that the various values  $s_i$  do not depend on  $\theta$  (or more precisely,  $p(s_1, s_2, \dots, s_n | \theta)$  does not involve  $\theta$ , in generalization of assumption 1) then he can specify that the  $m$ 's, given the  $s$ 's and  $\theta$  are multivariate normal, with mean and variances as in assumption 2 but with correlations  $\rho_{ij}$  between  $m_i$  and  $m_j$ . A new weighted average results: details are given in (A.2.1).

There is a difficulty here that is perhaps worthy of comment: namely how is N to make a reasoned judgment about the correlations? We offer two suggestions. Since correlation is a connection between only two subjects, it suffices to think of a pair of subjects,  $S_1$  and  $S_2$ . The distribution of  $m_1$ , given  $s_1$  and  $\theta$ , is normal with mean  $\alpha_1 + \beta_1 \theta$  and standard deviation  $\gamma_1 s_1$ . Consider next the distribution of  $m_2$ , given  $m_1$  (as well as  $s_1$ ,  $s_2$  and  $\theta$ ). In other words, suppose N had already received  $S_1$ 's judgment, what would he expect  $S_2$ 's to be? This requires a mean and a variance. Standard results show that, using a bivariate normal distribution for  $m_1$  and  $m_2$ , the latter is  $\gamma_2^2 s_2^2 (1 - \rho_{12}^2)$  instead of the  $\gamma_2^2 s_2^2$  had  $S_2$  been considered on his own. Thus  $1 - \rho_{12}^2$  is the proportionate reduction in the variance. A correlation of 1/2 reduces the variance by 25%. Hence the correlation may be interpreted as a variance reduction.

An alternative interpretation is through the mean of  $m_2$  given  $m_1$ . The normal distribution implies that this is equal to

$$\alpha_2 + \beta_2 \theta + (m_1 - \alpha_1 - \beta_1 \theta) \rho_{12} \gamma_2 s_2 / \gamma_1 s_1.$$

This has the form of the original mean plus a multiple of the deviation of  $m_1$  from its expectation; the multiple involving the correlation required. Consequently, the correlation has an alternative interpretation in the mean. Both interpretations should be used to give a check that the correlation is indeed of the magnitude suggested by either approach.

The following comment on independence and correlation may be appropriate. In a sense, subjects judging the same quantity are bound to be correlated merely because they are considering the same quantity. Thus if  $\theta$  is large, all the  $m_i$  will tend to be large. It is not this sort of correlation that is used in the argument presented here. That correlation, or independence, is for a given  $\theta$  and concerns the fluctuation of the  $m$ 's from that  $\theta$ . Thus we have to ask whether if  $S_1$  overestimates  $\theta$ ,  $m_1 > \theta$ , is  $S_2$  likely to do the same. Later (Section 6.0) similar considerations will apply to the scale information: if  $S_1$  thinks it is hard to estimate  $\theta$  ( $s_1$  large), will  $S_2$  think so too?

#### 4.2 Example

This section concludes with a numerical example adapted from Cohen and Brown (1980). Table 1 refers to three subjects. The first two columns give the mean range and standard deviation (in meters) for each of the subjects for a single target.  $S_1$  thought the target to be nearest but

TABLE 1

	m	s	$\alpha$	$\beta$	$\gamma$	
$s_1$	8000	2500	0	1.0	0.8	$\rho_{13}^2 = \frac{1}{2}$
$s_2$	14000	625	-500	1.1	1.2	
$s_3$	9650	900	-450	1.0	1.0	$\rho_{12} = \rho_{23} = 0$

was very uncertain.  $S_2$  was much more confident that it was far away, and  $S_3$  was intermediate both in range and error. The next three columns give N's assessments for the three subjects through the values of  $\alpha$ ,  $\beta$  and  $\gamma$ .  $S_1$  is thought to be without bias but to overestimate the error.  $S_2$  is thought to be without bias at 5,000 meters, but the bias increases by 10% of the distance over that amount; whereas, he tends to underestimate the error.  $S_3$  has a constant, negative bias but produces an estimate of error that N trusts. Finally,  $S_1$  and  $S_3$  are thought to be rather closely correlated with  $\rho_{13}^2 = \frac{1}{2}$  or  $\rho_{13} = 0.71$ ; so that, using the first interpretation above, the variance of  $S_3$  is reduced by one-half on N learning about  $S_1$ 's evaluation.  $S_2$  is a source independent of the other two. Table 2 shows in its first two columns the corrected values for each subject separately. Notice how the estimates and the errors are made more compatible than they were originally. The next three columns of Table 2 show the inverse of the variances and covariances, the elements  $\sigma^{ij}$  in the notation of the Appendix (A.2.1). Combination by (a.6) and (a.7) finally yield a mean range of 11,160 meters with a standard deviation of 540 meters.

TABLE 2

	$(m-\alpha)/\beta$	$\gamma s$	$s_1$	$s_2$	$s_3$
$s_1$	8000	2000	.333	0	-.370
$s_2$	13180	750	0	1.778	0
$s_3$	10100	900	-.370	0	1.646

## 5.0 t-DISTRIBUTIONS

It has been emphasized that the values of  $\alpha$ ,  $\beta$  and  $\gamma$  are known to N, though they might be improved by data from planned trials. It was suggested that estimates, like  $\hat{\gamma}$ , obtained from such trials could be used by N in subsequent evaluations using S's judgments. We return to the matter using a more refined analysis which has advantages in a quite different direction. For simplicity, attention will be confined to the case where S is known to have no bias,  $\alpha = 0$ ,  $\beta = 1$ , so that interest concentrates solely on  $\gamma$ .

What N has to provide is  $p(m|s, \theta)$ , or more fully  $p(m|s, \theta, \gamma)$ , since this distribution depends on  $\gamma$ . Let D denote the data obtained in trials of the sort described above, leading to N's opinion about  $\gamma$ ,  $p(\gamma|D)$ . Then  $\hat{\gamma}$  used above is the mode of this density. A complete analysis would not merely replace  $\gamma$  by  $\hat{\gamma}$  but would use the basic result that

$$p(m|s, \theta, D) = \int p(m|s, \theta, \gamma) p(\gamma|D) d\gamma. \quad (4.1)$$

In words, the probability for m, given s,  $\theta$  and D, is obtained by taking the same probability, given  $\gamma$  rather than D, and averaging it over the distribution of  $\gamma$  given D. It is shown in the Appendix (A.3) that the result of this calculation is that  $p(m|s, \theta, D)$  is no longer normal but has a t-distribution.

The original density  $p(m|s, \theta, \gamma)$  can be described by saying  $(m-\theta)/\gamma s$  has a normal distribution with zero mean and unit standard deviation. The new density  $p(m|s, \theta, D)$  is such that  $(m-\theta)/gs$  has a t-distribution with  $(n-1)$  degrees of freedom. Here

$$g^2 = \frac{1}{n-1} \sum \left( \frac{m_i - \theta}{s_i} \right)^2.$$

So what happens is that  $\gamma$  is replaced by  $g$  (which differs from the previous  $\hat{\gamma}$  only in using a divisor  $(n-1)$  instead of  $n$ ), and the normal distribution is replaced by a t-distribution. The latter differs from the former in assigning more probability to numerically large absolute values--remember both have mean zero. The difference is appreciable for small  $n$ , but rapidly diminishes for large  $n$ . Thus with  $n = 6$ , the familiar 2 standard deviations of the normal is replaced by 2.4. However, the use of  $t$  has an unexpected feature which is illustrated by an example.

Suppose there are two subjects.  $S_1$  has  $m_1 = -2.24$  and  $\gamma_1 s_1 = 1$ ,  $S_2$  has  $m_2 = +2.24$  and  $\gamma_2 s_2 = 1$ . If the distributions are normal, least squares will conclude that  $\theta$  has mean 0 and standard deviation 0.71.

(A weighted average of -2.24 and +2.24 with equal weights gives 0; assuming independence, the total precision is  $1+1=2$ , so the standard deviation is  $1/\sqrt{2} = 0.71$ .) Now this is somewhat unsatisfactory because, from  $S_1$ 's evidence,  $N$  thinks that  $\theta$  lies in the interval  $m_1$  plus or minus two standard deviations, that is  $(-4.24, -0.24)$ ; whereas on  $S_2$ 's evidence the interval is  $(+0.24, +4.24)$ ; and these two intervals do not overlap. Nevertheless,  $N$  averages the two to give  $(-1.42, +1.42)$  centered around 0, which both  $S_1$ 's and  $S_2$ 's evidence thinks is unlikely. What has happened is that least squares has forced a compromise between two conflicting opinions

and, worse still, says that the compromise value is more precise than either of the original ones. (Note that the compromise interval has width 2.84 against 4 for the two original intervals.)

If instead of N saying  $(m-\theta)/\gamma_s$  is normal, he says it is  $t$  with 5 d.f. then the 95% interval is  $(-4.81, 0.33)$  for  $S_1$  and  $(-0.33, +4.81)$  which do overlap slightly. Detailed calculations (outlined in the Appendix (A.4)) show the combined interval using both pieces of evidence is  $(-2.91, +2.91)$ . This has width 5.82 (compared with the normal value of 2.84) which exceeds the width, 5.14, of each of the original intervals. Consequently, the compromise centered at zero, which is compatible with both  $S_1$  and  $S_2$ , admits a degree of uncertainty over twice that of the normal compromise and greater than either subject separately suggests. Thus the  $t$ -distribution admits that the discrepancy between the subjects increases the uncertainty, whereas the normal distribution ignores the discrepancy and decreases the uncertainty.

Suppose the value 2.24 is increased, say to 3.0, so that  $S_1$  has  $m_1 = -3.0$ ,  $S_2$  has  $m_2 = +3.0$ , all the other values remaining unaltered while the discrepancy is increased. The normal result still persists and produces the combined interval of  $(-1.42, +1.42)$ . The  $t$ -distribution now refuses to admit a compromise, and the distribution for  $\theta$  given both pieces of evidence is bimodal with modes at  $\pm 2.0$ , and a local minimum at zero. We have not calculated the 95% interval but it will clearly be much wider than the earlier value 5.82.

The replacement of the normal by the t-distribution, therefore, has the great advantage that compromise is not forced nor is spurious accuracy claimed. It has the disadvantage over least-squares, based on the normal assumption, that no simple formulae are available for the results of combining t-distributions. Nevertheless, in these days of good calculators, the computation, once the program is written, takes only a matter of minutes.

Although the results using the t-distribution have been presented in the context of trials leading to values of  $g$  and the degrees of freedom  $v = n - 1$ , the distribution does not require this. Consider the following scenario. N was originally obliged to consider  $\gamma$  and found it difficult to fix on a value. After some thought he announced a value  $\gamma_0$  but said that it could be out by a factor of about 2. This factor is related to the value of  $v$ , if  $\gamma_0$  is used as  $g$ . We will see how to establish this relation below, but all we need note for the moment is that uncertainty about  $\gamma$  can be incorporated using  $g$  in place of  $\gamma$  and  $t$  for the normal,  $v$  incorporating the degree of uncertainty. The t-distribution is therefore of wide applicability, the additional quantity  $v$  providing substantial flexibility and realism.

## 6.0 SINGLE SUBJECT, WITH SCALE INFORMATION

The discussion so far has used assumption 1: that N judges that the standard deviation quoted by S, by itself, provides no information about  $\theta$ . We now relax this assumption. A possibility that will be investigated is that S's precision for  $\theta$  decreases with  $\theta$ , or that his standard deviation increases with  $\theta$ . In other words, N thinks that S will find large quantities harder to estimate than small ones. This is reasonably true in both the examples of target range and product demand. The simplest possibility is to suppose the standard deviation,  $s$ , increases linearly with  $\theta$ . This may be expressed by saying  $E(s|\theta) = \lambda\theta$ . For example, with  $\lambda = 0.1$ , N expects that S, faced with a target 10,000 meters away, will give a standard deviation of 1,000 meters; but at twice the distance, the deviation will be doubled. In addition, N will at least have to think about how much  $s$  might depart from its expectation. The easiest quantity to consider is the coefficient of variation of  $s$ . This is defined to be the ratio of the standard deviation (of  $s$ ) to the mean of  $s$ . And the simplest assumption is to suppose this not to change with  $\theta$ . We denote this constant by  $(\delta-1)^{-1}$ . (The notation is not deliberately awkward: a simple form here would lead to complications elsewhere.) For example, suppose  $\delta = 5$ , giving a coefficient of 0.5; since the mean in the above example was 1,000 meters, this leads to a standard deviation (of  $s$ ) of 500 meters.

A final consideration, beyond a mean of  $\lambda\theta$  and a coefficient of variation of  $(\delta-1)^{-\frac{1}{2}}$ , can enable N to fix his distribution for s, given  $\theta$ . It is unreasonable to suppose the distribution of s to be symmetric about the mean, a large increase being more likely than an equally large decrease. An increase from 1,000 meters to 2,000 (at two standard deviations) may be plausible, but a decrease to zero is not. This suggests using a distribution with a long tail to the right and a short one to the left, towards the origin. Such a distribution is obtained by supposing  $\lambda\delta\theta/s$  has a gamma distribution with parameter  $\delta$ : briefly, a  $\Gamma_{\delta}$  distribution. The Appendix (A.5) provides details. In applications it is probably easiest to think in terms of the logarithm of s,  $\ln s$ . This is approximately normal with mean  $\ln \lambda + \ln \theta$  and standard deviation  $(\delta-1)^{-\frac{1}{2}}$ . Returning to the numerical example with  $\lambda = 0.1$  and  $\theta = 5$ ,  $\ln s$  is about normal with mean  $\ln \theta = 2.30$  and standard deviation 0.25. For a target at 10,000 meters,  $\ln s$  has mean 6.91, so that the limits at two standard deviations are 6.41 and 7.41 which, translated back into meters, are 600 and 1,650 around the mean of 1,000. Thus N would be surprised if, with a target at 10,000 meters, S quoted a value for s more than 650 above or 400 below, the expected value of 1,000. This illustrates the lack of symmetry mentioned above.

The logarithmic transformation can also enable deviations by a factor to be handled. Thus at the end of Section 5.0 it was suggested that  $\gamma$  might differ from  $\gamma_0$  by a factor of 2; that is, limits for  $\gamma$  would be  $2\gamma_0$  and  $\frac{1}{2}\gamma_0$ . On taking logarithms the limits are  $\ln \gamma_0 \pm \ln 2$ , so that a standard deviation of  $\frac{1}{2}\ln 2 = 0.35$  is indicated.

We now have a distribution for  $s$ , given  $\theta$ . Retaining the normal distribution for  $m$ , given  $s$  and  $\theta$  (assumption 2) Bayes theorem (1) can be used to find  $N$ 's distribution for  $\theta$ , given  $m$  and  $s$ . The details are given in the Appendix (A.5). Essentially the result is a distribution for  $\theta$  of the form (a.9)

$$\propto \exp \left[ - \frac{(\theta - \mu)^2}{2\sigma^2} \right] \theta^{\delta+1} \quad (6.1)$$

where  $\propto$  is a constant,  $\delta$  is as before, and  $\mu$  and  $\sigma$  are quantities whose exact expressions in terms of  $m$  and  $s$  are given in the Appendix as

$$\mu = \frac{m-\alpha}{\beta} - \frac{\lambda\delta\gamma^2 s}{\beta^2} \quad \text{and} \quad \sigma = \gamma s \beta^{-1}.$$

#### 6.1 Approximations

The distribution (6.1) is of the form of a normal kernel, depending on  $\mu$  and  $\sigma$ , multiplied by a power of  $\theta$ . It is not an easy distribution to handle analytically but computation with it is straightforward and will be illustrated below. However, it is possible to obtain approximations to the mean and variance of  $\theta$ . We do not suggest that these are used in practice but they are presented here in order to help the reader understand the effect of the information provided by  $s$  alone. The results are obtained under the supposition that  $\sigma/\mu$  is small. In any case  $\sigma/\mu$  needs to be less than 3 for otherwise the normal kernel in (6.1) extends into negative values of  $\theta$ .

For the mean the approximation is (a.12)

$$E(\theta|m, s) = \frac{m-\alpha}{\beta} \left[ 1 + \frac{\gamma^2(\delta+1)s}{(m-\alpha)^2} \left( s - \frac{\hat{c}}{\delta+1} \frac{\lambda(m-\alpha)}{\beta} \right) \right].$$

To appreciate this recall that, with assumption 1 in which  $s$  alone gave no information about  $\theta$ , the mean was  $(m-\alpha)/\beta$ . The effect of the information provided by  $s$  is to multiply this value by something in square brackets a little different from one. Since  $\gamma^2(\delta+1)s/(m-\alpha)^2 > 0$ , this correction is above or below one according as  $s$  is greater than or less than  $\lambda(m-\alpha)/\beta$  times  $\delta/(\delta+1)$ . Since  $s$  was expected to be  $\lambda\theta$  and  $\theta$  is about  $(m-\alpha)/\beta$ , we see that, ignoring the factor  $\delta/(\delta+1)$ , the mean is increased from its value under assumption 1 if  $s$  exceeds its expectation and is otherwise decreased. (The factor  $\delta/(\delta+1)$  arises from the skewness of the distributions.) Reflection shows that this makes good sense.

For the variance the approximation is (a.11)

$$\text{var}(\theta|m, s) = \frac{\gamma^2 s^2}{\beta^2} \left[ 1 - (\delta+1) \frac{\gamma^2 s^2}{(m-\alpha)^2} \right]$$

which is always reduced from its value  $\gamma^2 s^2 / \beta^2$  when  $s$  alone contributes no information, the reduction being due to the extra information provided by  $s$ . Notice that the reduction increases with  $\delta$ , or increases as the coefficient of variation of  $s$ ,  $(\delta-1)^{-1/2}$ , decreases; again making good sense.

## 6.2 Example

We emphasize that these approximations are not very good except in extreme cases. So let us now turn to the exact calculations in a specific case.

N's judgment of a single subject S was expressed by the following values:

$$\alpha = 1,000; \beta = 1; \gamma = 1; \lambda = 0.1; \delta = 5.$$

Thus he thought S has a constant ( $\beta=1$ ) bias of 1,000 meters ( $\alpha$ ) but that his standard deviation was satisfactory ( $\gamma=1$ ). He expected a standard deviation of one tenth of the true range ( $\lambda=0.1$ ) and a coefficient of variation in it of  $\frac{1}{5}$ . (These are the numerical values of  $\lambda$  and  $\delta$  discussed above.)

The subject then quoted a range of  $m = 14,000$  meters and a standard deviation of  $s = 1,250$  meters. The values  $\mu$ ,  $\sigma$  and  $\delta + 1$  in (6.1) are easily seen to be

$$\mu = 12,375; \sigma = 1,250; \delta + 1 = 6.$$

The density for  $\theta$ , equation (6.1), is plotted in Figure 1 and labelled  $\lambda = 0.1$ . The mean,  $E(\theta|m, s)$ , is 13,100 meters and the standard deviation is 1,216 meters. Without the information provided by  $s$ , these values would have been 13,000 meters and 1,250 meters, respectively. The mean range has increased by 1% and the standard deviation has decreased by 3%. There is about 1 chance in 20 that  $\theta$  exceeds 14,900 meters and the same chance that it is less than 10,900 meters.

In this example, N expected  $s$  to be about one tenth ( $\lambda=0.1$ ) of  $\theta$ . Using  $(m-\alpha)/\beta$  for  $\theta$ , this is 13,000 and  $s$  is in reasonable agreement with this, at 1,250. To illustrate the effect of  $\lambda$  suppose it had been 0.2, so that N anticipated 2,600 rather than the low value he did receive. Notice that the low value is within reasonable limits for  $\theta$ . With  $s$  expected to be 2,600,  $\ln s$  is expected to be 7.863 with a standard

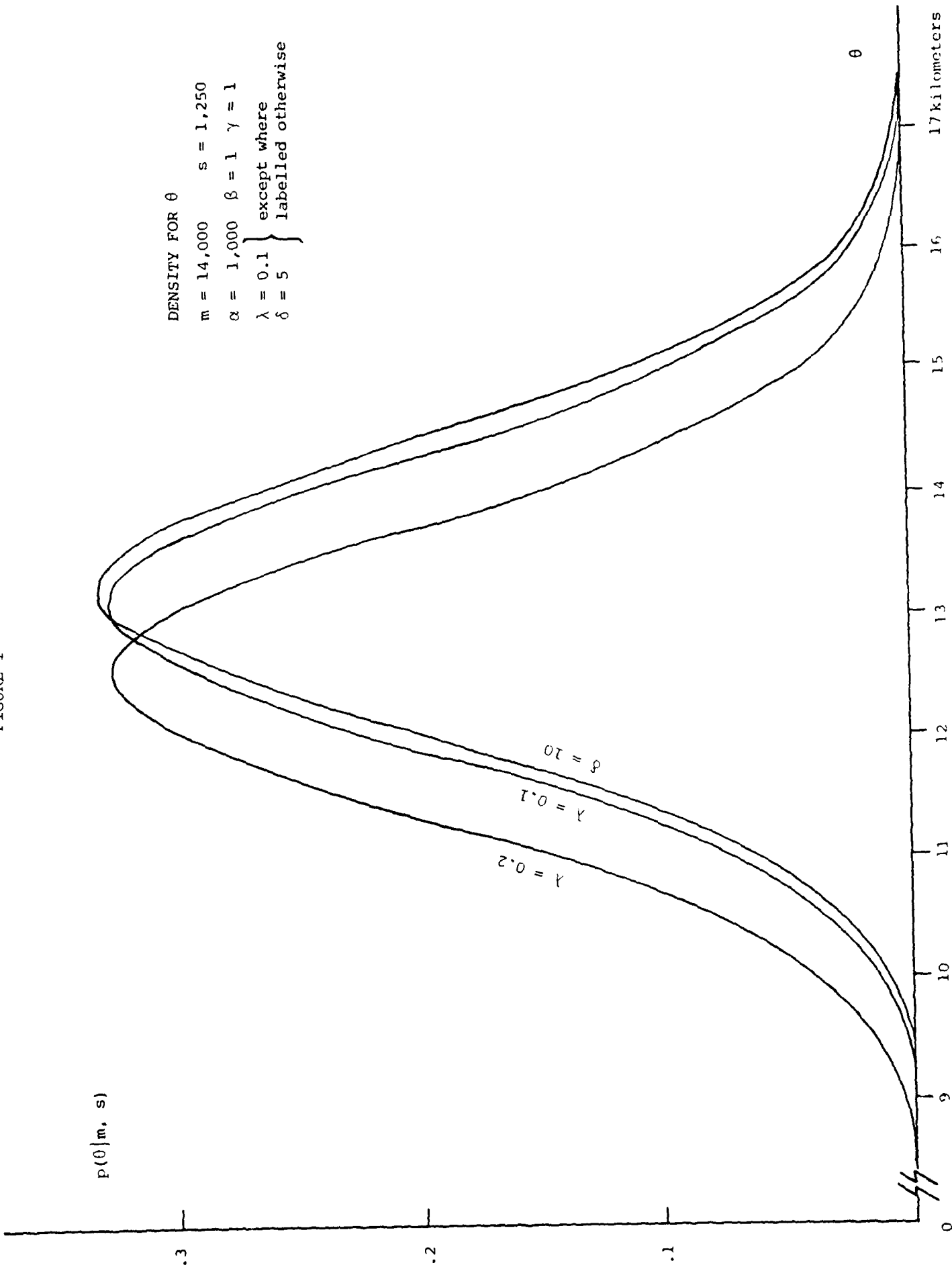
deviation of  $\frac{1}{3}$ . Two standard deviation limits are therefore 6.863 and 8.863, which, undoing the logarithm, give limits of 960 and 7,070 meters. Figure 1 graphs the density for  $\theta$ , where it is labelled  $\lambda = 0.2$ , having  $\mu = 11,750$ , but  $\sigma$  and  $\delta$  as before. The mean is 12,500 meters and the standard deviation 1,213 meters. The mean has decreased by 500 meters, or 4%, as a result of the unexpectedly low value of  $s$  but the standard deviation is unaltered. There is about 1 chance in 20 of  $\theta$  being above 14,200 meters or below 10,300 meters.

Finally suppose we revert to the original value of  $\lambda = 0.1$  but increase  $\delta$  to 10 so that  $N$  is more sure about the value of  $s$ , the coefficient of variation decreasing to  $1/3$ . Then

$$\mu = 11,875; \sigma = 1.250; \delta + 1 = 11.$$

The density is also plotted in Figure 1 and labelled  $\delta = 10$ . The mean is 13,190 meters with a standard deviation of 1,190 meters. The mean has increased by  $1\frac{1}{2}\%$  from its value without  $s$  and the standard deviation has decreased by about 5%. The latter reflects the stronger information provided by  $s$  due to the larger value of  $\delta$ . It is clear, however, that the effect of  $\delta$  is less than that of  $\lambda$ . This seems to be generally true.

FIGURE 1



## 7.0 SEVERAL SUBJECTS, SCALE INFORMATION

We now pass from the case of a single subject to several subjects. We saw above how to handle the  $m$ 's using a multivariate normal distribution with means  $\alpha_i + \beta_i \theta$  for  $S_i$ , variances  $\gamma_i^2 s_i^2$ , and correlations  $\rho_{ij}$ . The result was a normal distribution for  $\theta$ . Denote its mean by  $m_0$  and its standard deviation by  $s_0$ . (The formulae are given by (a.6) and (a.7) and were illustrated by a numerical example in Section 4.2.) We have also seen how to handle a single  $s_i$  by supposing  $\ln s_i$  is normal with mean  $\ln \gamma_i + \ln \theta$  and standard deviation  $(\delta_i - 1)^{-1/2}$ . If several subjects are involved, it can be assumed that the  $\ln s_i$  have a multivariate normal distribution, with means and standard deviations as just stated, and with correlations  $\tau_{ij}$ . As before, using (a.6) and (a.7) again, on the evidence of the  $s_i$ 's  $\ln \theta$  will have a normal distribution. Let its mean be written  $\ln \lambda_0$  and its standard deviation  $\delta_0^{-1/2}$ . Then reversing the connection between a  $\Gamma_{\delta}$ -distribution and the normal, this implies that to a good approximation  $\theta \delta_0 / \lambda_0$  has a  $\Gamma_{\delta_0 - 1}$  distribution. This may be combined with the normal density for  $\theta$  obtained from the means with the result that the final distribution of  $\theta$  is still of the form (6.1) above. The detailed formulae are given in the Appendix (A.7).

## 8.0 DISCUSSION

The description of a subject's ability as a probability assessor through the parameters  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\lambda$ , and  $\epsilon$  allows a considerable amount of flexibility, yet never leads to distributions more complicated than a normal kernel times a power (equation (6.1)). This distribution persists even when correlations are admitted between the stated means, and the stated variances, of different subjects. Although approximations are available, exact computations can easily be performed on a simple calculator and, previously programmed, are both simple and fast to use.

The parameters describe one's knowledge of the subject. Unfortunately, in practical situations this knowledge is not extensive and the parameters depend almost entirely on personal impressions. These have their place and are often of great importance but, nevertheless, we need much more experience in carefully controlled trials of subjects' ability as assessors. A lot of work in this field is concerned with untrained subjects, naive in their knowledge of probability. The trials needed are with highly skilled people, like sonarmen, who have had probability training. Such trials would also help to determine whether the assumptions of normality, gamma, or  $t$ , made in this paper, are realistic descriptions of subjects' performance.

#### REFERENCES

Bordley, R.F., and Wolff, R.W. On the aggregation of individual probability estimates. Management Science, 1981, 27, 959-964.

Cohen, M.S., and Brown, R.V. Decision support for attack submarine commanders (Technical Report 80-11). Falls Church, VA: Decision Science Consortium, Inc., 1980.

DeGroot, M.H. Improving judgments using feedback. Trab. Estadist, 1980, 31, 383-395.

French, S. Updating of belief in the light of someone else's opinion. Journal of the Royal Statistical Society, Series A, 1980, 143, 43-48.

Lindley, D.V. The improvement of probability judgments. Journal of the Royal Statistical Society, Series A, 1982, 145 (to appear).

Lindley, D.V., Tversky, A., and Brown, R.V. On the reconciliation of probability assessments. Journal of the Royal Statistical Society, Series A, 1979, 142, 146-180.

Morris, P.A. Decision analysis expert use. Management Science, 1974, 20, 1233-1241.

Morris, P.A. Combining expert judgments: A Bayesian approach. Management Science, 1977, 23, 679-693.

## APPENDIX

### A.1 Single Subject, No Scale Information

Assumption 1.  $p(s|\theta)$  does not involve  $\theta$ .

Assumption 2.  $p(m|s, \theta)$  is  $N(\alpha + \beta\theta, \gamma s)$ .

Then

$$p(\theta|m, s) \propto \exp \left\{ -\frac{1}{2} \left( \frac{m - \alpha - \beta\theta}{\gamma s} \right)^2 \right\} p(\theta) \quad (a.1)$$

and if  $p(\theta)$  is uniform over the effective range of the likelihood,

$$p(\theta|m, s) \propto \exp \left\{ -\frac{1}{2} \left[ \frac{\theta - (m - \alpha)/\beta}{\gamma \beta^{-1} s} \right]^2 \right\} \quad (a.2)$$

which is  $N((m - \alpha)/\beta, \gamma \beta^{-1} s)$

A.1.1 Subject trials. If  $\beta = 1$ , and  $n$  trials give  $(m_i, s_i, \theta_i)$  then the likelihood for  $\alpha$  and  $\gamma$  is

$$\gamma^{-n} \exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left( \frac{m_i - \alpha - \theta_i}{\gamma s_i} \right)^2 \right\}$$

and the maximum likelihood estimates are easily seen to be

$$\left. \begin{aligned} \hat{\alpha} &= \sum (m_i - \theta_i) s_i^{-2} / \sum s_i^{-2} \\ \text{and} \\ \hat{\gamma}^2 &= n^{-1} \sum (m_i - \theta_i - \hat{\alpha})^2 / s_i^2 \end{aligned} \right\} \quad (a.3)$$

## A.2 Several Subjects, No Scale Information

For several subjects, judged independent, (a.1) gives a likelihood for  $\theta$

$$\exp \left\{ -\frac{1}{2} \sum_{i=1}^n \left( \frac{m_i - \alpha_i - \beta_i \theta}{\gamma_i s_i} \right)^2 \right\},$$

giving another normal distribution with mean

$$\left\{ \sum \beta_i (m_i - \alpha_i) / \gamma_i^2 s_i^2 \right\} / \left\{ \sum \beta_i^2 / \gamma_i^2 s_i^2 \right\} \quad (\text{a.4})$$

and variance

$$\left\{ \sum \beta_i^2 / \gamma_i^2 s_i^2 \right\}^{-1}. \quad (\text{a.5})$$

(a.4) is the usual weighted average of the individual values  $(m_i - \alpha_i) / \beta_i$  with weights inversely proportional to the variances  $\gamma_i^2 s_i^2 / \beta_i^2$ ; and (a.5) adds up the individual precisions (inverses of variances).

A.2.1 Correlations between subjects. If correlations between the subjects are included by replacing the separate normal distributions for each  $m_i$ , given  $s_i$  and  $\theta$ , by a multivariate normal distribution, the means and variances will remain the same at  $\alpha_i + \beta_i \theta$  and  $\gamma_i s_i$ , but there will be a covariance  $\rho_{ij} \gamma_i \gamma_j s_i s_j$ , yielding a correlation  $\rho_{ij}$  between  $m_i$  and  $m_j$ . For ease of notation write  $c_{ij}$  for the covariance, equal to the variance when  $i = j$ , and let  $\sigma^{ij}$  denote the elements of the matrix inverse to the variance-covariance matrix with elements  $c_{ij}$ . The multivariate normal distribution gives a likelihood for  $\theta$

$$\begin{aligned} & \exp \left[ -\frac{1}{2} \sum (m_i - \alpha_i - \beta_i \theta) \sigma^{ij} (m_j - \alpha_j - \beta_j \theta) \right] \\ & \propto \exp \left[ -\frac{1}{2} \left\{ \theta^2 \sum \beta_i \sigma^{ij} \beta_j - 2\theta \sum \beta_i \sigma^{ij} (m_j - \alpha_j) \right\} \right] \end{aligned}$$

$$\propto \exp \left[ -\frac{1}{2} \sum_i \beta_i \sigma^{ij} \beta_j \left\{ \theta - \frac{\sum_i \beta_i \sigma^{ij} (m_j - \alpha_j)}{\sum_i \beta_i \sigma^{ij} \beta_j} \right\}^2 \right]$$

where all summations are over both  $i$  and  $j$ . The mean is

$$\sum_i \beta_i \sigma^{ij} (m_j - \alpha_j) / \sum_i \beta_i \sigma^{ij} \beta_j \quad (\text{a.6})$$

and the variance

$$(\sum_i \beta_i \sigma^{ij} \beta_j)^{-1}. \quad (\text{a.7})$$

The mean is still a weighted average of the individual contributions

$$(m_j - \alpha_j) / \beta_j \text{ with weights } \sum_i \beta_i \sigma^{ij} \beta_j.$$

### A.3 Uncertainty About $\gamma$ : t-Distribution

With  $\alpha = 0$ ,  $\beta = 1$  in  $n$  trials yielding  $(m_i, s_i, \theta_i)$ , the likelihood for  $\gamma$  is (see A.1.1)

$$\gamma^{-n} \exp \left\{ -\frac{1}{2\gamma^2} \sum_{i=1}^n (m_i - \theta_i)^2 / s_i^2 \right\}$$

and, assuming a uniform distribution for  $\gamma$  before the trials, this is proportional to  $p(\gamma|D)$ . If we write  $g^2 = \frac{1}{n-1} \sum \left( \frac{m_i - \theta_i}{s_i} \right)^2$  we have from (4.1)

$$p(m|s, \theta, D) \propto \int \exp \left[ -\frac{(m-\theta)^2}{2\gamma^2 s^2} - \frac{(n-1)g^2}{2\gamma^2} \right] \gamma^{-(n+1)} d\gamma$$

and the standard gamma integral shows this to be proportional to

$$(1 + t^2/v)^{-\frac{1}{2}(v+1)} \text{ where}$$

$$t = \frac{m-\theta}{gs}, \quad \text{and} \quad v = n-1$$

so that the quantity called  $t$  has a  $t$ -distribution with  $v$  degrees of freedom.

#### A.4 Products of t-Likelihoods

Suppose that N's judgment is that the subjects are independent, assumption 1 holds, and that for  $S_i$ ,  $(m_i - \alpha_i - \theta) / \beta_i g_i s_i$  has a t-distribution with  $v_i$  degrees of freedom, replacing the normal distribution previously assumed. Then the likelihood for  $\theta$  from all subjects is the product of the separate t-likelihoods

$$\prod_i \left\{ 1 + \frac{(m_i - \alpha_i - \theta)^2}{v_i \beta_i^2 g_i^2 s_i^2} \right\}^{-\frac{1}{2}(v_i + 1)}$$

Such products are not easy to handle analytically though computation is straightforward. A simple example will illustrate the principle phenomena.

Consider two t-distributions with the same degrees of freedom  $v = v_1 = v_2$ ;  $\alpha_i = 0$ ,  $\beta_i = 1$  (no biases);  $g_1 s_1 = g_2 s_2 = s$ , say (equal precision); with one centered at  $m_1 = +m$  and the other at  $m_2 = -m$ . The likelihood is then

$$\left\{ 1 + \frac{(\theta - m)^2}{s^2 v} \right\} \left\{ 1 + \frac{(\theta + m)^2}{s^2 v} \right\} \quad (a.8)$$

raised to the power  $-\frac{1}{2}(v+1)$ . If  $\phi = \theta/sv^{\frac{1}{2}}$  and  $c = m/sv^{\frac{1}{2}}$ , a little algebra reduces (a.8) to

$$1 + \left( \frac{\phi^2 + 1 - c^2}{2c} \right)^2$$

apart from a constant multiple. The resulting density in  $\phi$  is therefore bimodal if  $c^2 > 1$ , the modes occurring at  $\phi = \pm (c^2 - 1)^{\frac{1}{2}}$ . If  $c^2 \leq 1$ , there is a single mode at  $\phi = 0$ . The example given in the text had  $g_1 s_1 = s = 1$ ,

$v = 5$ , and  $m = 2.24 = \sqrt{5}$ , so that  $c = 1$ . Later  $m$  was increased to 3.0 giving  $c = 1.34$ , and two modes. The interval for  $\theta$  quoted was evaluated using numerical integration.

#### A.5 Single Subject, Scale Information

Assumption 3.  $p(s|\theta)$  is such that  $\lambda\delta\theta/s$  has a  $\Gamma_{\delta}$  distribution.

$X$  has a  $\Gamma_{\delta}$  distribution if its density is, for  $X \geq 0$ ,

$$e^{-X} X^{\delta} / \delta!$$

with  $\delta > -1$ . It is usual to write  $X = \frac{1}{2}\chi^2$  and  $\delta = \frac{1}{2}v - 1$ , and refer to the  $\chi^2$ -distribution with  $v$  degrees of freedom; but the  $\frac{1}{2}$ 's that occur are a nuisance in the present context. The density for  $s$  is then

$$p(s|\theta) = \exp\{-\lambda\delta\theta/s\} (\lambda\delta\theta)^{\delta+1} / s^{\delta+2} \delta!$$

and simple calculations show that  $E(s) = \lambda\theta$  and  $\text{var}(s) = \lambda^2\theta^2/(\delta+1)$ . The coefficient of variation is therefore  $(\delta+1)^{-1/2}$ . The logarithm of  $s$ ,  $\ln s$ , is approximately normal with mean  $\ln(\lambda\theta) - \frac{1}{2}(\delta+1)^{-1/2}$  and standard deviation  $(\delta+1)^{-1/2}$ . (In the moment results it is being assumed that  $\delta > 1$ .)  $\frac{1}{2}(\delta+1)^{-1/2}$  is typically small in comparison with  $\ln \lambda\theta$  and will be omitted in the analyses.

Retaining assumption 2 and using Bayes theorem

$$\begin{aligned} p(\theta|m, s) &\propto \exp \left[ -\frac{1}{2} \left( \frac{m-\alpha-\beta\theta}{\gamma s} \right)^2 - \frac{\lambda\delta\theta}{s} \right] \theta^{\delta+1} p(\theta) \\ &\propto \exp \left[ -\frac{\beta^2}{2\gamma^2 s^2} \left( \theta - \frac{E(m-\alpha)-\lambda\delta\gamma^2 s}{\beta^2} \right)^2 \right] \theta^{\delta+1} \end{aligned}$$

if  $p(\theta)$  is locally constant. The form of the distribution of  $\theta$  is that of a normal kernel multiplied by  $\theta^{\delta+1}$ . Write  $\mu$  and  $\sigma$  for the mean and standard deviation respectively of the normal kernel (not of  $\theta$ ): that is, write

$$\mu = [\delta(m-\alpha) - \lambda\delta\gamma^2s] \theta^{-2}$$

and

$$\sigma = \gamma s \theta^{-1},$$

so that

$$p(\theta|m, s) \propto \exp \left[ - \frac{(\theta - \mu)^2}{2\sigma^2} \right] \theta^{\delta+1}. \quad (a.9)$$

**A.5.1 Approximations.** To study this distribution it is necessary to evaluate the missing constant of proportionality. Writing  $\theta^{\delta+1} = (\theta - \mu + \mu)^{\delta+1}$  and expanding by the binomial theorem the integral of (a.9) is easily obtained as a finite series in terms of the known moments about the mean for a normal distribution. Moments may similarly be found since  $E(\theta^r)$  leads to the same integral with  $\delta$  replaced by  $\delta + r$ . Since the  $r$ th moment for a normal distribution is proportional to  $\sigma^r$ , the series will be in terms of  $\sigma/\mu = \tau$ . Consequently, if the coefficient of variation,  $\sigma/\mu$ , of the normal kernel is small, reasonable approximations to the moments can be obtained by including only the first few terms of the series. Straightforward, but tedious, algebra shows that

$$\left. \begin{aligned} \text{and} \quad E(\theta|m, s) &= \mu \{ 1 + (\delta+1)\tau^2 + O(\tau^4) \} \\ \text{var}(\theta|m, s) &= \sigma^2 \{ 1 - (\delta+1)\tau^2 + O(\tau^4) \} \end{aligned} \right\} \quad (a.10)$$

The approximations obtained by omitting the  $O(\tau^4)$  terms will only be

reasonable if  $(\delta+1)\tau^2$  is small. Inserting the values of  $\mu$  and  $\sigma$  into these approximation yields

$$E(\theta|m, s) = \left( \frac{m-\alpha}{\beta} - \frac{\lambda\delta\gamma^2 s}{\beta^2} \right) \left( 1 + (\delta+1) \frac{\gamma^2 s^2}{(m-\alpha)^2} \right)$$

where, in the correction  $(\delta+1)\tau^2$ ,  $\mu$  has been replaced by  $(m-\alpha)/\beta$  since the correction to this value will overall result in a higher-order correction, and

$$\text{var}(\theta|m, s) = \frac{\gamma^2 s^2}{\beta^2} \left( 1 - (\delta+1) \frac{\gamma^2 s^2}{(m-\alpha)^2} \right). \quad (\text{a.11})$$

The expression for the mean can be simplified since the product of the two terms involving  $\gamma^2$  can be neglected in comparison with  $O(\tau^4)$  already discarded. Hence

$$E(\theta|m, s) = \frac{m-\alpha}{\beta} \left[ 1 + \frac{\gamma^2 (\delta+1) s}{(m-\alpha)^2} \left( s - \frac{\beta}{\delta+1} \frac{\lambda(m-\alpha)}{\beta} \right) \right] \quad (\text{a.12})$$

#### A.6 Other Forms of Scale Information

We next make some comments on assumption 3. There are other possibilities: for example  $\lambda^2 \delta \theta^2 / s^2$  may be supposed to have a  $\Gamma_\delta$  distribution. This is equivalent to working in terms of the variance  $s^2$  rather than the standard deviation as assumption 3 does. The revised assumption implies that  $E(s) = \lambda\theta$ , as before, except that this is only approximate. ( $E(s) = \lambda\theta(1 - \frac{1}{8}\delta)$  is a better approximation.) The coefficient of variation is approximately  $2\delta^{-\frac{1}{2}}$ , as against  $(\delta-1)^{-\frac{1}{2}}$  with assumption 3. Calculations with this assumption yield results in close agreement with those reported above provided the  $\delta$  of the new method,  $\delta_1$  say, is related to that of the

original by  $2\delta_1^{-\frac{1}{2}} = (\delta-1)^{-\frac{1}{2}}$  to preserve the equality of the coefficients of variation.

Both these assumptions imply that  $s$  increases linearly with  $\theta$ . Another possibility is that the variance  $s^2$  increases linearly with  $\theta$ . In analogy with assumption 3, it can be supposed that  $\lambda\delta\theta/s^2$  is  $\Gamma_\delta$ ; or in line with the ideas in the last paragraph  $\lambda\delta\theta^2/s^4$  is  $\Gamma_\delta$ . Calculations show that the effect on the original normal distribution, obtained under assumption 1 of  $s$  alone giving no information about  $\theta$ , is slight. This is because the standard deviation is about linear with  $\theta^{\frac{1}{2}}$ ; in other words, the rate of change of  $s$  with  $\theta$  is so slow that no appreciable disturbance from normality is felt.

#### A.7 Several Subjects, Scale Information

With several subjects suppose, as described in the text, that on the evidence of the means,  $\theta$  is normal with mean  $m_0$  and variance  $s_0^2$ , whereas on the evidence of the standard deviation alone it is such that  $\theta\delta_0/\lambda_0$  is  $\Gamma_{\delta_0-1}$ . Then the final distribution is the product of these, namely,

$$\begin{aligned} & \exp \left\{ -\frac{1}{2} \left( \frac{\theta - m_0}{s_0} \right)^2 - \frac{\theta\delta_0}{\lambda_0} \right\} \theta^{\delta_0-1} \\ & \propto \exp \left\{ -\frac{1}{2s_0^2} \left( \theta - m_0 + \frac{\delta_0 s_0^2}{\lambda_0} \right)^2 \right\} e^{\delta_0-1} \end{aligned}$$

of the same form as studied earlier.

DISTRIBUTION LIST

Office of Naval Research

Stuart L. Brodsky  
Group Leader/Mathematics  
Code 411MA  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Neal D. Glassman  
Code 411MA  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Charles J. Holland  
Code 411MA  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

CDR Kent Hull  
Code 410B  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Mr. Gerald Malecki  
Engineering Psychology Programs  
Code 442  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Charles R. Paoletti  
Administrative Contracting Officer  
Room 718  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Dr. James C. T. Pool  
Division Leader/Mathematical  
& Physical Sciences  
Code 411  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

J. Randolph Simpson (10 copies)  
Scientific Officer, Code 431  
Naval Analysis Programs  
Mathematical and Information  
Sciences Division  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Dr. Martin A. Tolcott  
Division Leader, Psychological  
Sciences Division  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Dr. Thomas C. Varley  
Group Leader  
Code 411OR  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Edward J. Wegman  
Group Leader, Code 411S&P  
Office of Naval Research  
800 N. Quincy Street  
Arlington, VA 22217

Office of Naval Research  
Eastern/Central Regional Office  
666 Summer St.  
Boston, MA 02210

DISTRIBUTION LIST (continued)

Other Government Agencies

Defense Technical Information  
Center (12 copies)  
Cameron Station  
Building 5  
Alexandria, VA 22314

Naval Research Laboratory (6 copies)  
Attn: Code 2627  
Washington, D.C. 20375

Supplemental Distribution

Professor Kenneth J. Arrow  
Department of Economics  
Stanford University  
Stanford, CA 94305

Michael Athens  
Massachusetts Institute of Technology  
77 Massachusetts Avenue  
Cambridge, MA 02139

Dr. David Castanon  
Alpha-Tech  
Burlington, MA 01803

Dr. Ward Edwards  
Director  
Social Sciences Research Institute  
1042 W. 36th Street  
Denney Research Building  
Los Angeles, CA 90007

Supplemental Distribution (cont.)

Dr. Peter Farquhar II  
Graduate School of Administration  
Voorhies Hall  
University of California  
Davis, CA 95616

Irwin R. Goodman  
Naval Ocean Systems Center  
Code 7232  
San Diego, CA 92152

Professor James March  
Graduate School of Business  
Administration  
Stanford University  
Stanford, CA 94305

Dr. Amos Tversky  
Department of Psychology  
Stanford University  
Stanford, CA 94305

Professor Robert Wilson  
Department of Business Administration  
Stanford University  
Stanford, CA 94305

Professor L. Zadeh  
Department of Engineering  
University of California  
Berkeley, CA 94701

**DATE**  
**ILME**